

Strongly self-interacting processes on the circle

Carl-Erik Gauthier* and Pierre Monmarché†

June 10, 2016

Abstract

The purpose of this paper is to investigate the long time behaviour for a self-interacting diffusion and a self-interacting velocity jump process. While the diffusion case has already been studied for some particular potential function, the second one, which belongs to the family of piecewise deterministic processes, is new.

Depending on the underlying potential function's shape, we prove either the almost sure convergence or the recurrence for a natural extended process given by a change a variable.

Keywords: Self-Interacting Markov processes, diffusions, PDMP, ergodicity, almost-sure convergence.

MSC primary: 60K35, 60J25, 60H10, 60J75, 60J60

1 Introduction

Our aim is to study the effect of the addition of a self-interaction mechanism to two initially Markovian dynamics. The first one is the classical Fokker-Planck diffusion $X \in \mathbb{R}$ that solves the SDE

$$dX_t = dB_t - V'(X_t)dt,$$

where $(B_t)_{t \geq 0}$ is a standard Brownian motion on \mathbb{R} . Namely X is the Markov process with generator

$$Lf(x) = \frac{1}{2}f''(x) - V'(x)f'(x).$$

We recall the generator of a Markov process $(Z_t)_{t \geq 0}$ is formally defined by

$$Lf(z) = (\partial_t)_{|t=0} \mathbb{E}(f(Z_t) \mid Z_0 = z).$$

*Institut de Mathématiques, Université de Neuchâtel, Rue Émile Argand 11, 2000 Neuchâtel, Switzerland. Email: carl-erik.gauthier@unine.ch

†Institut de Mathématiques, Université de Neuchâtel, Rue Émile Argand 11, 2000 Neuchâtel, Switzerland. Email: pierre.monmarche@unine.ch

The second one is the velocity jump process $(X, Y) \in \mathbb{R} \times \{-1, 1\}$ which is the piecewise deterministic Markov process (PDMP) introduced in [14] with generator

$$Lf(x, y) = y\partial_x f(x, y) + (\lambda + (yV'(x))_+) (f(x, -y) - f(x, y))$$

where $\lambda > 0$ is constant and $(\cdot)_+$ denotes the positive part (see [14] and Section 2.2 for a trajectory definition of the dynamic). In both cases, if we suppose that the potential V is sufficiently coercive at infinity, X is ergodic and its law converges to the Gibbs measure with density proportional to e^{-V} . Note that when the rate of jump λ goes to infinity and time is correctly accelerated, the velocity Gibbs process (more precisely its first coordinate) converges to the Fokker-Planck diffusion (see [7]).

In both cases we want to replace the potential $V(X_t)$ by a self-interacting potential

$$V_t(X_t) = \int_0^t W(X_t, X_s) ds$$

where W is a symmetric interaction potential. In other words $V_t(X_t)$ depends both on the current position X_t and the (non-normalized) occupation measure $\int_0^t \delta_{X_s} ds$. This is a strong self-interaction, by contrast with the weak self-interaction such as studied in [2] where the self-interacting potential is a function of X_t and of the normalized occupation measure $\frac{1}{t} \int_0^t \delta_{X_s} ds$.

Self-Interacting processes belong to the family of *path-dependent* processes. The particularity of such processes is their lack of Markov property since the past modifies the environment that drives the particle. New phenomena may arise in their long time behaviour, which would be impossible without the path-dependency.

A first example of strong self-interaction is the linear one, that correspond to $W(x, y) = \frac{1}{2}(x - y)^2$. M.Cranston and Y.Le Jan proved in 1995 (see [6]) the almost sure convergence of the solution of the SDE

$$dX_t = dB_t - \int_0^t (X_t - X_s) ds. \quad (1)$$

Later, S.Herrmann and B.Roynette extended this result to a broader class $W(x, y) = V(x - y)$ with V convex (see [10]). In the case of the circle, the first author obtained the same result for $W(x, y) = -\cos(x - y)$ (see [8]). In all these cases the particle is attracted by its past.

In [1], M.Benaïm and the first author considered the repulsive case, in which the particle is repelled by its past trajectory. More precisely they studied a self-repelling diffusion on a compact manifold where W can be decomposed as

$$W(x, y) = \sum_{i=1}^n a_i e_i(x) e_i(y)$$

with the a_i 's being positive numbers and the e_i 's being eigenfunctions of the Laplace operator on the manifold. The basic example on the circle would be $W(x, y) = \cos(x - y) = \cos(x)\cos(y) + \sin(x)\sin(y)$. This assumption on the e_i 's yields an explicit formula for the invariant measure of the Markov process $\left(X_t, \left(\int_0^t e_i(X_s) ds\right)_{i=1..n}\right)$.

The aim of the present work is to investigate the case where the e_i 's are not eigenfunctions of the Laplace operator. On the other hand we restrict the study (in dimension 1) to the case $n = 1$, namely we take a potential of the form

$$W(x, y) = F(x)F(y)$$

with moreover F smooth and 2π -periodic, so that we consider $x \in \mathbb{S}^1 = \mathbb{R}/2\pi\mathbb{Z}$. Following [1], we set

$$U_t = \int_0^t F(X_s) ds, \quad (2)$$

which reduces the study of the non-Markovian process to the study of some Markov process on an extended space. This restriction should be seen as a first step toward the analysis of the more general situation.

As a consequence, in this paper we study the Markov processes (X, U) on $\mathbb{S}^1 \times \mathbb{R}$ and (X, U, Y) on $\mathbb{S}^1 \times \mathbb{R} \times \{-1, 1\}$ with respective generators

$$L_1 f(x, u) = \frac{1}{2} \partial_x^2 f(x, u) - u F'(x) \partial_x f(x, u) + F(x) \partial_u f(x, u) \quad (3)$$

and

$$L_2 f(x, u, y) = y \partial_x f(x, u, y) + F(x) \partial_u f(x, u, y) + (\lambda + (yu F'(x))_+) (f(x, u, -y) - f(x, u, y)). \quad (4)$$

In both cases we call X the position, U the auxiliary variable and, in the case of the velocity jump process, Y the velocity. We work under the following assumption:

- The function $F : \mathbb{S}^1 \rightarrow \mathbb{R}$ is non-constant, smooth, changes signs, and $F'(x) = 0$ implies $F(x) \neq 0$. Moreover for all $x \in \mathbb{S}^1$ there exists $k \geq 1$ such that $F^{(k)}(x) \neq 0$. In particular the critical points of F are isolated points.

Throughout this paper, we consider the discrete sets

$$\begin{aligned} M(F, +) &= \{x \in \mathbb{S}^1 \mid x \text{ is a local maximum of } F \text{ and } F(x) > 0\} \\ M(F, -) &= \{x \in \mathbb{S}^1 \mid x \text{ is a local maximum of } F \text{ and } F(x) < 0\} \\ m(F, +) &= \{x \in \mathbb{S}^1 \mid x \text{ is a local minimum of } F \text{ and } F(x) > 0\} \\ m(F, -) &= \{x \in \mathbb{S}^1 \mid x \text{ is a local minimum of } F \text{ and } F(x) < 0\} \end{aligned}$$

and $\mathcal{M} = M(F, -) \cup m(F, +)$. Recall the total variation distance between two probability laws μ and ν is

$$d_{TV}(\mu, \nu) = \inf \{ \mathbb{P}(\Xi_1 \neq \Xi_2), \text{ Law}(\Xi_1) = \mu, \text{ Law}(\Xi_2) = \nu \}$$

and a measure μ is said invariant for a Markov process $(Z_t)_{t \geq 0}$ if $\{\text{Law}(Z_0) = \mu\}$ implies $\{\forall t \geq 0, \text{Law}(Z_t) = \mu\}$. We say that the law of $(Z_t)_{t \geq 0}$ converges exponentially fast to μ in the total variation sense if there exist $C, \rho > 0$, that may depend on the law of Z_0 , such that

$$d_{TV}(\text{Law}(Z_t), \mu) \leq C e^{-\rho t}.$$

Our main result is the following:

Theorem 1.

1. If $\mathcal{M} = \emptyset$, then each of the processes (X, U) with generator (3) and (X, U, Y) with generator (4) admits a unique invariant measure. If the law of U_0 admits an exponential moment then the process converges exponentially fast in the total variation distance sense to this invariant measure.
2. If $\mathcal{M} \neq \emptyset$, then, in both cases, the position X_t almost surely converges as t goes to infinity to a point of \mathcal{M} . Any point of \mathcal{M} has a positive probability to be the limit of X .

Before proceeding to its proof, let us mention why this result may be expected. Suppose that, at some time, $U > 0$. Then, as long as U is large enough, the force $U_t F'(X_t)$ tends to confine X close to the minima of F . If these minima are all negative, while X stays in their neighbourhood, U decreases, up to some point where it becomes negative. From then the effect of the force is reversed, X is attracted by the maxima of F , and the same mechanism comes into play with U and F changed to $-U$ and $-F$. In some sense X and U have then an inhibitory effect one on the other.

On the other hand if X falls in the neighbourhood of a positive minimum of F while $U > 0$ (the case of a negative maximum with $U < 0$ being symmetric) then, as long as it stays there, U increases, which make it more and more unlikely for X to escape away from the minimum, so that eventually there is a positive probability that X never leaves and U goes to infinity. This is reminiscent of the annealing problem (see [16] for the diffusion and [14] for the velocity jump process) where U_t is replaced by a deterministic $(\beta_t)_{t \geq 0}$, called the inverse temperature. It is classical that in this case, if β increases faster than logarithmically then X will eventually stay trapped forever in the cusp of a local minima. Yet, in our present case, as long as X stays close to a positive minimum, U increases linearly in time.

Remark 1.

1. The particular form of the interacting potential $W(x, y) = F(x)F(y)$ implies that W is a Mercer Kernel, which means the particle is repulsed by its past (see [1]). We could also consider the case $W(x, y) = -F(x)F(y)$. Following the proof of Theorem 1, it is not hard to see that in this case X_t almost surely converges as t goes to infinity to a point of $\mathcal{M}' = m(F, -) \cup M(F, +)$ which, as soon as F is not constant and changes signs, is non-empty.
2. If F does not change signs, then, depending on the sign, U_t converges either to ∞ or to $-\infty$ linearly fast. Therefore, Proposition 1 and Proposition 4 imply the almost-sure convergence of X_t respectively either to a local minimum or to a local maximum of F .

We made the choice to write as much as possible notations, results and proofs which are common to both processes, isolating only the few lemmas that deal with the specific technical difficulties of each case. Our arguments are based on bounds for some hitting

times of the processes which are established in Section 2. From them we show in Section 3 that, when \mathcal{M} is empty, the time for the processes to return to compact sets is short (i.e. in a time with exponential moments). Section 4 is devoted to some uniform bounds of the transition kernel of the processes over compact sets, and Section 5 to the proof of Theorem 1.

2 Hitting times

In this section, for a redaction purpose, we will hide the dependency on U of the evolution of X . More precisely we will consider the (inhomogeneous in time) diffusion

$$dX_t = dB_t - g(t)F'(X_t)dt \quad (5)$$

for any Lipschitz function g and, similarly, the inhomogeneous PDMP (X, Y) with generator

$$L_t f(x, y) = y \partial_x f(x, y) + (\lambda + (g(t)yF'(x))_+) (f(x, -y) - f(x, y)) \quad (6)$$

where the generator of an inhomogeneous Markov process Z is by definition

$$L_t f(z) = (\partial_s)_{|s=0} \mathbb{E}(f(Z_{t+s}) \mid Z_t = z).$$

Note the processes considered in Theorem 1 are particular cases of those defined here.

Let $A = m(F, +) \cup m(F, -)$ be the set of minima of F , and $\delta \leq -\frac{1}{3} \max\{F(x) : x \in m(F, -)\}$ be positive and small enough so that

- for all $x \in A$, denoting by $I_x^\delta = [z_l, z_r]$ the connected component of $\{F \leq F(x) + 2\delta\}$ containing x , then F decreases on $[z_l, x]$ and increases on $[x, z_r]$.
- there exists $\kappa > 0$ such that for all $x \in A$ and $\eta \in [0, \delta]$,

$$d(x, B_x^\eta) \geq \kappa \sqrt{\eta},$$

where $B_x^\eta = \{z \in I_x^\delta, F(z) = F(x) + \eta\}$.

Finally, let

$$B^\eta = \bigcup_{x \in A} B_x^\eta \quad \text{and} \quad C^\eta = \left(\bigcup_{x \in A} I_x^\eta \right)^c.$$

In other words C^η is the complementary of a neighbourhood of the minima of F and B^η is a set of intermediary points from A to C^η . These sets (for $\eta = \delta$) are represented in Fig. 1. Note that the choice of δ ensures that if $\mathcal{M} = \emptyset$ then C^δ contains $\{F \geq -\delta\}$.

For $x \in \mathbb{S}^1$ and $D \subset \mathbb{S}^1$ we write

$$\begin{aligned} T_{x \rightarrow D} &= \inf \{t \leq 0, X_t \in D \mid X_0 = x\} \\ q_{x \rightarrow D} &= \mathbb{P}(X \text{ reaches } D \text{ before } A \mid X_0 = x) \end{aligned}$$

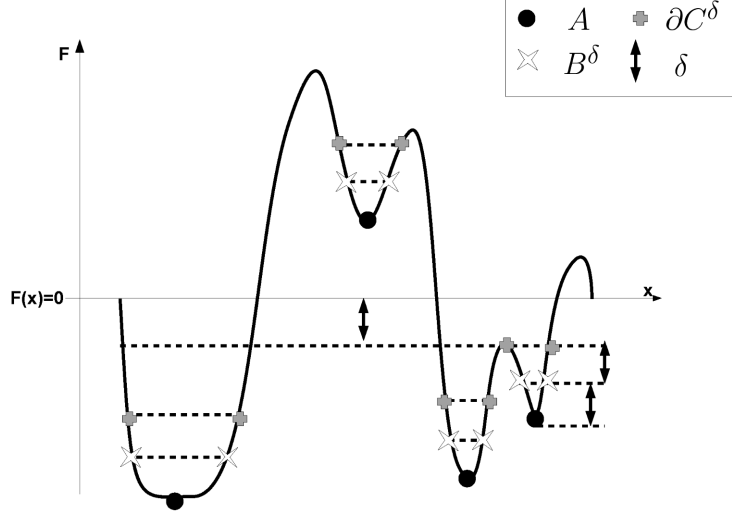


Figure 1: Starting from a minimum in A , the process has to cross an intermediary point of B^δ halfway before reaching C^δ . The energy level difference from A to B^δ , from B^δ to C^δ or (for a negative minimum) from ∂C^δ to $\{F = 0\}$ is always at least δ .

For two real random variables V, W , recall that V is said to be stochastically smaller than W , denoted by $V \leq^{sto} W$ (or equivalently $W \geq^{sto} V$), if for all $r \in \mathbb{R}$

$$\mathbb{P}(V > r) \leq \mathbb{P}(W > r).$$

If V and W have same law we write $V \stackrel{law}{=} W$.

The aim of this section is to prove the following:

Proposition 1. *There exist a constant $K > 1$ and nonnegatives random variables S and R with some finite positive exponential moments, such that for all $M > 1$ and $\eta \in (0, \delta]$ with $M\eta > 1$, for all Lipschitz function $g \geq M$, if X is defined by (5) or if (X, Y) is defined by the generator (6), the following holds:*

$$\forall x \in B^\eta, \quad q_{x \rightarrow C^\eta} \leq K M e^{-\eta M} \quad (7)$$

$$\forall x \in A, \quad T_{x \rightarrow B^\eta} \stackrel{sto}{\geq} R \sqrt{\eta}. \quad (8)$$

$$\forall x \in \mathbb{S}^1, \quad T_{x \rightarrow A} \stackrel{sto}{\leq} S \quad (9)$$

Remark 2. *In the case of the velocity jump process (X, Y) , note that these bounds are uniform over the initial velocity Y_0 .*

The meaning of these bounds is the following. Suppose the auxiliary variable U (whose role here is played by an arbitrary function g) stays for some time above a given level $M > 1$. Then the position X will fall in a local minima of F within a time shorter than S , which does not depend on M (i.e. a high U can only accelerate the hitting time of A). Then to climb back up to an intermediary point of B^η , it takes a time $R\sqrt{\eta}$, which

is again uniform on $M > 1$. From B^η , the probability to escape from the neighbourhood of the minimum in one attempt (namely to reach C^η before having fallen back to A , the bottom) is of order $e^{-\eta M}$, which is a classical metastability result (see [4, 14] for instance) if g is thought as an inverse temperature, since η is the potential barrier to overcome.

The proof of Proposition 1 is split in the two next subsections since the arguments are different for each dynamic. Note that in several proofs we will make assumptions like $x_1 \leq x_2$ where x_1 and x_2 are in \mathbb{S}^1 , which will make sense since at these times we will only be concerned by the behaviour of the processes on given simply connected intervals of \mathbb{S}^1 .

2.1 For the diffusion

Proof of Inequality (7) in the diffusion case. Consider the diffusion defined by (5) with $g \geq M$ and $X_0 = x \in B^\eta$. Since $g \geq M$, it follows from Ikeda-Watanabe's comparison result [11, Theorem 1.1, Chapter VI] that

$$q_{x \rightarrow C^\eta} \leq \mathbb{P} \left(\tilde{X} \text{ hits } C^\eta \text{ before } A \mid \tilde{X}_0 = x \right) := \tilde{q}_{x \rightarrow C^\eta},$$

where \tilde{X} solves the SDE

$$d\tilde{X}_t = dB_t - MF'(\tilde{X}_t) dt.$$

Its scale function is defined by

$$\begin{aligned} p(y) &= \int_x^y \exp(-2 \int_x^z -MF'(s) ds) dz \\ &= \int_x^y e^{2M(F(z)-F(x))} dz. \end{aligned}$$

Let $x_0 \in A$ and $x_1 \in C^\eta$ be such that F is monotonous on the interval between x_0 and x_1 that contains x . Suppose without loss of generality that $x_0 < x < x_1$. By [12, Proposition 5.22, Chapter 5.5],

$$\tilde{q}_{x \rightarrow C^\eta} = \frac{p(x_1) - p(x)}{p(x_1) - p(x_0)} \leq \frac{2\pi e^{2M\eta}}{\int_{x_0}^{x_1} e^{2M(F(z)-F(x))} dz}$$

where we used the local monotonicity of F . On the other hand,

$$\begin{aligned} \int_{x_0}^{x_1} e^{2M(F(z)-F(x))} dz &\geq \frac{1}{2M\|F'\|_\infty} \int_{x_0}^{x_1} 2MF'(z) e^{2M(F(z)-F(x))} dz \\ &= \frac{1}{2M\|F'\|_\infty} (e^{4M\eta} - 1). \end{aligned}$$

Therefore, as $M\eta > 1$,

$$q_{x \rightarrow C^\eta} \leq 4\pi M\|F'\|_\infty \frac{e^{2M\eta}}{e^{4M\eta} - 1} \leq 8\pi M\|F'\|_\infty e^{-2M\eta}.$$

□

To prove the two other assertions of Proposition 1, we need the following comparison result:

Lemma 1. *Let x_0 be a local extrema of F and $\varepsilon > 0$ be such that F' is monotonous on $J_\varepsilon := (x_0 - \varepsilon, x_0 + \varepsilon)$. Consider X the diffusion defined by (5), with $g \geq M > 1$, starting at $X_0 = x \in J_\varepsilon$, and W a standard Brownian motion. Denote by*

$$\chi^\varepsilon(x) = \inf \{t > 0, X_t \notin J_\varepsilon\} \quad \text{and} \quad \iota^\varepsilon = \inf \{t > 0, |W_t| = \varepsilon\}$$

the respective exit time from J_ε of X and $x_0 + W$. Then:

1. *if x_0 is a local maximum of F , for all $x \in J_\varepsilon$,*

$$\chi^\varepsilon(x) \stackrel{sto}{\leq} \iota^\varepsilon.$$

2. *if x_0 is a local minimum of F ,*

$$\chi^\varepsilon(x_0) \stackrel{sto}{\geq} \iota^\varepsilon.$$

Proof. First, note that by symmetry the exit time from J_ε of $x + W$ has the same law as the exit time of $x + 2(x_0 - x) + W$, and since the process $x_0 + W$ necessarily crosses x or $x + 2(x_0 - x)$ before leaving J_ε , the exit time of $x_0 + W$ is stochastically greater than the one of $x + W$ for any $x \in J_\varepsilon$.

Consider $\Theta = (X - x_0)^2$, which solves

$$d\Theta_t = 2\sqrt{\Theta_t}d\tilde{B}_t + dt + 2g(t)((X_t - x_0)F'(X_t))dt,$$

where $\tilde{B}_t = \int_0^t \text{sign}(X_s - x_0)dB_s$ is still a standard Brownian motion. Then $(X_0 - x_0 + \tilde{B})^2$ is a weak solution of

$$dZ_t = 2\sqrt{Z_t}d\tilde{B}_t + dt.$$

When x_0 is a maximum (resp. minimum) of F , $(x - x_0)F'(x)$ is non-positive (resp. non-negative) on J_ε , so that by Ikeda-Watanabe's comparison result, $\Theta_t \geq Z_t$ (resp. $\Theta_t \leq Z_t$) up to the first time where Θ reaches ε^2 . As a conclusion, when x_0 is a maximum, Θ reaches ε^2 before Z , and thus in a time stochastically greater than ι^ε , and when x_0 is a minimum, Θ reaches ε^2 after Z and the latter happens at a time with law ι^ε if the starting point is x_0 . \square

Proof of Inequality (8) in the diffusion case. Recall that there exists a constant $\kappa > 0$ such that for all $x \in A$ and $\eta < \delta$, $d(x, B_x^\eta) \geq \kappa\sqrt{\eta}$. From Lemma 1 and the Brownian motion's scaling property,

$$T_{x \rightarrow B^\eta} \geq \chi(\kappa\sqrt{\eta}, x) \stackrel{sto}{\geq} \iota^{\kappa\sqrt{\eta}} \stackrel{law}{=} \eta^{\frac{1}{4}} \iota^\kappa.$$

The fact that ι^κ has an exponential moment is a consequence of [5, Theorem 2]. \square

Proof of Inequality (9) in the diffusion case. For a given small enough $\varepsilon > 0$, denote by

$$E^\varepsilon = \bigcup_{x \in M(F,+) \cup M(F,-)} (x - \varepsilon, x + \varepsilon)$$

the set of points which are at a distance less than ε from a maximum of F . Let X be the diffusion defined by (5) with $g \geq M$. We apply the following procedure:

1. If, at some time, $X_t \in E^\varepsilon$, wait until it leaves E^ε , which according to the first part of Lemma 1 happens in a time stochastically smaller than ι^ε .
2. If at some time t_0 , X leaves E^ε , compare it with $X_{t_0} + B$ where B is the Brownian motion that drives the SDE (5). More precisely by Ikeda-Watanabe's comparison result, $F(X_t) \leq F(X_{t_0} + B_t)$ up to the time where either X or B reach an extremum of F .
3. Wait until B reaches an extrema of F . If this is a maximum, go back to the first step. If this is a minimum then necessarily, at this time, X has already crossed this minimum, stop the procedure.

Note that, ε being fixed, the probability that $x_0 + B$ reaches a maximum rather than a minimum is bounded above by some $p < 1$ which is uniform over all $x_0 \in \partial E_\varepsilon$. Hence the number of iteration of the procedure is stochastically less than a geometric random variable G with parameter p . Conditionally to whether the Brownian motion reaches a minimum or a maximum in step 3, the law of the duration of the third step is different, but in either cases it is stochastically smaller than $\iota^{2\pi}$. Therefore the total duration of one iteration of the procedure is stochastically smaller than ι^C for some constant $C > 0$, independently from whether this is the last iteration or not. Let $(\iota_k)_{k \geq 0}$ be i.i.d copies of ι^C , independent from G .

We have obtained that for all $x \in \mathbb{S}^1$,

$$T_{x \rightarrow A} \stackrel{sto}{\leq} \sum_{k=0}^G \iota_k$$

so that

$$\mathbb{E} \left(e^{cT_{x \rightarrow A}} \right) \leq \mathbb{E} \left(\left(\mathbb{E} (e^{c\iota_0}) \right)^G \right)$$

which is finite for c small enough. □

2.2 For the velocity jump process

This subsection is devoted to the proof Proposition 1 in the PDMP case, namely for the inhomogeneous Markov process (X, Y) with generator (6). First we construct a trajectory of the process (X, Y) in the following way: consider two independent i.i.d. sequences of standard (with mean 1) exponential random variables $(E_i)_{i \in \mathbb{N}}$ and $(F_i)_{i \in \mathbb{N}}$. Set $T_0 = 0$ and

suppose the process has been defined up to some time T_k independently from $(E_i, F_i)_{i \geq k}$. Let

$$\begin{aligned}\theta_1 &= \inf \left\{ t > 0, \int_0^t g(T_k + s)(Y_{T_k} F'(X_{T_k} + sY_{T_k}))_+ ds > E_k \right\}, \\ \theta_2 &= \frac{1}{\lambda} F_k,\end{aligned}$$

and $T_{k+1} = T_k + \theta_1 \wedge \theta_2$, which is the next jump time. If $T_{k+1} = T_k + \theta_1$ we say that the jump is due to the landscape, else we say it is due to the constant rate λ . In either cases, set $X_t = X_{T_k} + (t - T_k)Y_{T_k}$ for all $t \in [T_k, T_{k+1}]$, $Y_t = Y_{T_k}$ for all $t \in [T_k, T_{k+1})$ and $Y_{T_{k+1}} = -Y_{T_k}$. Thus by induction the process is defined up to time T_n for all n . Note that even if, depending on g , the rate of jump may not be bounded, two jumps due to the landscape cannot be arbitrarily close (since at such a jump time, $yF'(x)$ becomes non-positive), so that there cannot be infinitely many jumps in a finite time and $T_n \rightarrow \infty$ as $n \rightarrow \infty$.

Proof of Inequality (7) in the PDMP case. We mainly have to adapt to our inhomogeneous settings the proof of [14, Proposition 4.1]. Without loss of generality, we consider the following configuration: $x_0 \in A$, $x_1 \in B^\eta$ and $x_2 \in \partial C^\eta$ with $x_0 < x_1 < x_2$, and F is increasing on $[x_0, x_2]$.

Let $M \geq 1$ and \mathcal{L}_M be the set of Lipschitz functions $g \geq M$. For all $x \in [x_1, x_2]$, set

$$\eta_x = \sup_{g \in \mathcal{L}_M} \mathbb{P}((X, Y) \text{ reaches } (x_2, 1) \text{ before } (x, -1) \mid (X_0, Y_0) = (x, 1)).$$

where the supremum runs over the function g that appears in the generator (6) of the process (X, Y) .

Consider a process (X, Y) with generator (6) with some function $g \in \mathcal{L}_M$. For a small $\varepsilon > 0$, suppose that $(X_0, Y_0) = (x - \varepsilon, 1)$.

Then the probability that X goes from $x - \varepsilon$ to x without any jump is less than $1 - \varepsilon(MF'(x) + \lambda) + o_{\varepsilon \rightarrow 0}(\varepsilon)$ and the probability it reaches $(x, 1)$ before $(x - \varepsilon, -1)$ but with at least one jump is of order ε^2 as $\varepsilon \rightarrow 0$.

If the process has reached $(x, 1)$, it has a probability less than η_x to reach $(x_2, 1)$ before having fallen back to $(x, -1)$. Nevertheless, if indeed it has fallen back to $(x, -1)$, it has a probability $\varepsilon\lambda + o_{\varepsilon \rightarrow 0}(\varepsilon)$ to jump before reaching $(x - \varepsilon, -1)$, in which case it reaches again $(x, 1)$ with probability $1 + o_{\varepsilon \rightarrow 0}(1)$. In this latter case, it reaches $(x_2, 1)$ before $(x - \varepsilon, -1)$ with probability less than $\eta_x + o_{\varepsilon \rightarrow 0}(1)$. Thus everything boils down to

$$\begin{aligned}\eta_{x-\varepsilon} &\leq (1 - \varepsilon(MF'(x) + \lambda)) \eta_x (1 + \varepsilon\lambda) + o_{\varepsilon \rightarrow 0}(\varepsilon) \\ &= (1 - \varepsilon MF'(x)) \eta_x + o_{\varepsilon \rightarrow 0}(\varepsilon).\end{aligned}$$

Together with $\eta_{x_2} = 1$, it yields $\eta_x \leq e^{-M(F(x_2) - F(x))}$, and in particular $\eta_{x_1} \leq e^{-\eta M}$.

Let

$$r_y = \sup_{g \in \mathcal{L}_M} \mathbb{P}((X, Y) \text{ reaches } (x_2, 1) \text{ before } (x_0, -1) \mid (X_0, Y_0) = (x_1, y)).$$

Starting from $(x_1, -1)$ and until the process either jumps or reaches $(x_0, -1)$, we have $YF'(X) < 0$ so that, whatever the function g in (6) is, there cannot be any jump due to the landscape during this time. On the other hand if $\theta_2 > 2\pi$, which happens with probability $e^{-2\lambda\pi}$, there is also no jump due to the constant rate during this time, so that

$$\mathbb{P}((X, Y) \text{ reaches } (x_1, 1) \text{ before } (x_0, -1) \mid (X_0, Y_0) = (x_1, -1)) \leq 1 - e^{-2\lambda\pi}.$$

On the one hand it means that $r_{-1} \leq (1 - e^{-2\lambda\pi}) r_1$ and on the other hand that

$$r_1 \leq \eta_{x_1} + (1 - e^{-2\lambda\pi}) r_1$$

and finally that

$$q_{x_1 \rightarrow C^\eta} \leq \max(r_1, r_{-1}) \leq e^{2\lambda\pi} e^{-\eta M}.$$

□

Proof of Inequality (8) in the PDMP case. Since $|Y| = 1$, the time needed to reach B^η from A is deterministically larger than $d(A, B^\eta) \geq \kappa\sqrt{\eta}$. □

Proof of Inequality (9) in the PDMP case. Suppose that, at some point in the construction of a trajectory, $\theta_2 > 4\pi$, which happens with probability $e^{-4\lambda\pi}$. If there is also no jump due to the landscape in the meanwhile, X covers the whole circle and in particular reaches A in a time less than 2π . On the other hand if there is a jump due to the landscape before time 2π , the velocity turns to its opposite, and from then and up to the hitting time of A , $YF'(X) < 0$, so that in the meanwhile there cannot be another jump due to the landscape: A is attained in a time less than 4π .

It means that as soon as $\theta_2 > 4\pi$, X reaches A in a time less than 4π , so that starting from any point of \mathbb{S}^1 , X reaches A in a time stochastically smaller than $4\pi G$ where G is a geometric variable with parameter $e^{-4\lambda\pi}$. □

3 Stability

In this section we consider either $Z = (X, U)$ or $Z = (X, U, Y)$ such as in Theorem 1, and we are interested in the time of return of Z to compact sets. More precisely for $M > 1$ we write

$$\tau_M = \inf\{t > 0, |U_t| \leq M\}$$

and we want to prove τ_M admits exponential moments. The constant K and the random variables R, S appearing along this section are those given by Proposition 1.

Lemma 2. *Suppose $\mathcal{M} = \emptyset$. Let $M > 1$ be such that $KMe^{-\delta M} < 1$, and let $(S_i)_{i \in \mathbb{N}}$, $(R_i)_{i \in \mathbb{N}}$ and $(G_i)_{i \in \mathbb{N}}$ be independent i.i.d. sequences where S_0 (resp. R_0) is a copy of S (resp. $\sqrt{\delta}R$) and G_0 has geometric law with parameter $KMe^{-\delta M}$. For $t \geq 0$ let*

$$N_t = \inf \left\{ n \in \mathbb{N}, \sum_{k=1}^{G_0 + \dots + G_n} R_k \geq t \right\}.$$

Then for all $t > 0$ and for any initial condition Z_0 with $U_0 > M$,

$$\int_0^{t \wedge \tau_M} \mathbb{1}_{\{F(X_s) \geq -\delta\}} ds \stackrel{sto}{\leq} \sum_{k=0}^{N_t} S_k.$$

Proof. While $t \leq \tau_M$, the estimates of Proposition 1 hold for X . In particular, independently from its initial condition, the process reaches A in a time stochastically smaller than S_0 . Then it takes at least a time R_1 to climb back to B^δ . From there, it reaches C^δ with probability less than $KMe^{-\delta M}$, else it falls back to A . Therefore it remains a time stochastically greater than $\sum_{k=1}^{G_0} R_k$ in $(C^\delta)^c = \{F \leq -\delta\}$ before reaching C^δ . When this finally occurs, the process falls again back to A after a time less than S_1 (independently from what occurred before it had reached C). We call this an excursion in C^δ . After n excursions, the process has stayed at least a time $\sum_{k=1}^{G_0 + \dots + G_n} R_k$ in $\{F \leq -\delta\}$, which implies in particular that at time t there have been stochastically less than N_t excursions. Thus during a time t , the time spent in C^δ is stochastically less than $\sum_{k=0}^{N_t} S_k$. \square

Recall that from Cramer's Theorem (see e.g [15, Chapter 2.4] with the exercise 2.28 in it), if $(V_i)_{i \geq 0}$ is an i.i.d. sequence of non-negative variables with some finite positive exponential moments, then $(\frac{1}{n} \sum_{i=0}^n V_i)_{n \geq 0}$ satisfies a Large Deviation Principle, in the sense there exists $c_1, c_2 > 0$ such that for all $n \geq 0$,

$$\mathbb{P} \left(\left| \frac{1}{n} \sum_{i=0}^n V_i - \mathbb{E}(V) \right| > \frac{1}{2} \right) \leq c_1 e^{-c_2 n}.$$

Proposition 2. Suppose $\mathcal{M} = \emptyset$. Then for $M > 0$ large enough there exist $\zeta > 0$ and $C < \infty$ such that for all initial condition z_0

$$\mathbb{E}_{z_0} (e^{\zeta \tau_M}) \leq e^{C\zeta|u_0|}.$$

Proof. We only treat the case $u_0 > M > 1$ (the case $u_0 < -M$ is obtained by changing U and F to their opposite). Moreover we suppose $KMe^{-\delta M} < 1$. From

$$U_t \leq u_0 - \delta t + (\delta + \max F) \int_0^t \mathbb{1}_{\{F(X_s) \geq -\delta\}} ds$$

we deduce

$$\begin{aligned} \mathbb{P}(\tau_M > t) &= \mathbb{P} \left(\tau_M > t \text{ and } \int_0^t \mathbb{1}_{\{F(X_s) \geq -\delta\}} ds > \frac{M - u_0 + \delta t}{\delta + \max F} \right) \\ &\leq \mathbb{P} \left(\sum_{k=0}^{N_t} S_k > \frac{M - u_0 + \delta t}{\delta + \max F} \right) \end{aligned}$$

where N_t is defined in Lemma 2. For any $a, b \in \mathbb{N}$,

$$\left\{ \sum_{i=0}^b G_i \geq a \quad \text{and} \quad \sum_{i=1}^a R_i \geq t \right\} \subset \{N_t \leq b\}$$

and

$$\left\{ N_t \leq b \quad \text{and} \quad \sum_{i=0}^b S_i \leq \frac{M - U_0 + \delta t}{\delta + \max F} \right\} \subset \left\{ \sum_{i=0}^{N_t} S_i \leq \frac{M - u_0 + \delta t}{\delta + \max F} \right\},$$

which implies

$$\mathbb{P}(\tau_M > t) \leq \mathbb{P} \left(\sum_{i=0}^b G_i < a \quad \text{or} \quad \sum_{i=1}^a R_i < t \quad \text{or} \quad \sum_{i=0}^b S_i > \frac{M - u_0 + \delta t}{\delta + \max F} \right).$$

Applied with $a = a_t = \lceil \frac{2t}{\mathbb{E}(R_1)} \rceil$ and $b = b_t = \lceil \frac{2a_t}{\mathbb{E}(G_1)} \rceil$, it implies

$$\begin{aligned} \mathbb{P}(\tau_M > t) &\leq \mathbb{P} \left(\frac{1}{b_t} \sum_{i=1}^{b_t} \frac{G_i}{\mathbb{E}(G_i)} < \frac{1}{2} \right) + \mathbb{P} \left(\frac{1}{a_t} \sum_{i=1}^{a_t} \frac{R_i}{\mathbb{E}(R_i)} < \frac{1}{2} \right) \\ &\quad + \mathbb{P} \left(\sum_{i=0}^{b_t} S_i > \frac{M - u_0 + \delta t}{\delta + \max F} \right). \end{aligned}$$

For $t \geq \mathbb{E}(R_1)$, we have

$$b_t \leq \left(2KM e^{-\delta M} \left(\frac{2t}{\mathbb{E}(R_1)} + 1 \right) + 1 \right) \leq \frac{7KM e^{-\delta M}}{\mathbb{E}(R_1)} t.$$

Therefore, for $t \geq 1 + \mathbb{E}(R_1) + 2\frac{1}{\delta}u_0$ and M large enough,

$$\begin{aligned} \mathbb{P} \left(\sum_{i=0}^{b_t} S_i > \frac{M - u_0 + \delta t}{\delta + \max F} \right) &\leq \mathbb{P} \left(\frac{1}{b_t} \sum_{i=0}^{b_t} S_i > \frac{\delta t}{4(\delta + \max F)b_t} \right) \\ &\leq \mathbb{P} \left(\frac{1}{b_t} \sum_{i=0}^{b_t} S_i > \frac{\delta \mathbb{E}(R_1) e^{\delta M}}{28KM(\delta + \max F)} \right) \\ &\leq \mathbb{P} \left(\frac{1}{b_t} \sum_{i=0}^{b_t} \frac{S_i}{\mathbb{E}(S_i)} > 2 \right). \end{aligned}$$

For such large M and t , the Large Deviation Principle satisfied by R, S and G_0 implies

$$\mathbb{P}(\tau_M > t) \leq \beta e^{-\rho t}$$

for some $\beta, \rho > 0$ which do not depend on u_0 . The proof is completed with

$$\mathbb{E}(e^{\zeta \tau_M}) = 1 + \zeta \int_0^\infty e^{\zeta s} \mathbb{P}(\tau_M > s) ds.$$

□

Remark 3. The statement of the proposition remains valid for any initial condition z_0 such that $u_0 > 0$ (resp. $u_0 < 0$) under the weaker assumption $m(F, +) = \emptyset$ (resp. $M(F, -) = \emptyset$).

4 Transition kernel bounds

In this section we still consider either $Z = (X, U)$ or $Z = (X, U, Y)$ such as in Theorem 1, and we call E its state space, namely either $\mathbb{S}^1 \times \mathbb{R}_+$ or $\mathbb{S}^1 \times \mathbb{R}_+ \times \{-1, 1\}$. We aim to prove the following local Doeblin condition holds:

Proposition 3. *Let \mathcal{K} be a compact set of E . There exist $t_0 > 0$, $0 < c < 1$ and a probability measure ν on E such that for all $z \in \mathcal{K}$, for all Borel set D ,*

$$\mathbb{P}(Z_{t_0} \in D \mid Z_0 = z) \geq c\nu(D).$$

For the diffusion process, this classically follows from an hypoellipticity argument. By contrast, note that the velocity jump process is not regularizing, in the sense its transition kernel is never absolutely continuous with respect to the Lebesgue measure (at all time there is a positive probability that the process hasn't jumped yet). However the Doeblin condition can still be obtained from some controllability property and a partial regularization.

Since, again, the arguments are different for both processes, we split the proof of Proposition 3 in two paragraphs.

4.1 For the diffusion

In this subsection we consider the process $Z = (X, U)$ induced by the generator (3), namely the solution of the SDE

$$\begin{cases} dX_t &= dB_t - U_t F'(X_t) dt \\ dU_t &= F(X_t) dt. \end{cases} \quad (10)$$

Lemma 3. *For all $z_0 \in \mathbb{S}^1 \times \mathbb{R}$ and $t > 0$, the transition kernel $\mathbb{P}(Z_t \in \cdot \mid Z_0 = z_0)$ admits a smooth density with respect to the Lebesgue measure and its support is $\mathbb{S}^1 \times [u_0 + (\min F)t, u_0 + (\max F)t]$.*

Proof. For $(x, u) \in \mathbb{S}^1 \times \mathbb{R}$, set

$$G_0(x, u) = \begin{pmatrix} 1 \\ 0 \end{pmatrix} \text{ and } G_1(x, u) = \begin{pmatrix} -uF'(x) \\ F(x) \end{pmatrix}.$$

The Lie-bracket of G_0, G_1 is the vector field $[G_0, G_1]$ given by

$$[G_0, G_1](x, u) = DG_1(x, u)G_0(x, u) - DG_0(x, u)G_1(x, u) = \partial_x G_1(x, u) = \begin{pmatrix} -uF''(x) \\ F'(x) \end{pmatrix}.$$

So by iteration, we have

$$\underbrace{[G_0, [G_0, \dots [G_0, G_1] \dots]]}_{k \text{ times}}(x, u) = \partial_x^{(k)} G_1(x, u) = \begin{pmatrix} -uF^{(k+1)}(x) \\ F^{(k)}(x) \end{pmatrix}.$$

Therefore, by our non-degeneracy assumption on F , the SDE (10) satisfies everywhere the Hörmander condition (see for instance [9]), which gives the first part of the proposition. For the second part, first note that for $z = (x_0, u_0) \in \mathbb{S}^1 \times \mathbb{R}$,

$$\left((X_t, U_t) \right)_{t \geq 0} \subset \mathbb{S}^1 \times [u_0 + (\min F)t, u_0 + (\max F)t].$$

Now let $((x_s, u_s))_{s \geq 0}$ denotes the solution of the ordinary differential equation

$$\begin{cases} \dot{x} &= v(t) - uF'(x) \\ \dot{u} &= F(x) \end{cases}$$

with initial condition $(x(0), u(0)) = z$ and where $s \mapsto v(s)$ is a piecewise constant function. Given $z' = (x', u') \in \mathbb{S}^1 \times (u_0 + (\min F)t, u_0 + (\max F)t)$, we aim to build a function v such that $(x(t), u(t))$ is arbitrarily close to z' . Let $\varepsilon \in (0, t)$ be arbitrary small and $t_0, t_1 \geq 0$ such that $t_0 + t_1 = t$ and $u' - u_0 = t_0(\min F) + t_1(\max F)$.

First, choose $v_0 \in \mathbb{R}$ such that $v(s) = v_0$ for all $s \in [0, \varepsilon]$ and $x(\varepsilon) \in \{y \in \mathbb{S}^1 \text{ s.t. } F(y) = \min F\}$ and let $v(s) = 0$ for $s \in (\varepsilon, t_0]$. Then, pick $v_1 \in \mathbb{R}$ such that $v(s) = v_1$ for all $s \in (t_0, t_0 + \varepsilon]$ and $x(t_0 + \varepsilon) \in \{y \in \mathbb{S}^1 \text{ s.t. } F(y) = \max F\}$ and let $v(s) = 0$ for $s \in (t_0 + \varepsilon, t - \varepsilon]$. Finally, choose $v_2 \in \mathbb{R}$ such that $v(s) = v_2$ for all $s \in (t - \varepsilon, t]$ and $x(t) = x'$.

Note that $u(t) = u' + o_{\varepsilon \rightarrow 0}(1)$. The Stroock-Varadhan support's Theorem concludes. \square

Proof of Proposition 3 in the diffusion case. Denoting by $p_t(\cdot, \cdot)$ the transition density given by Lemma 3, let $z_1, z_2 \in E$ be such that $p_{t_1}(z_1, z_2) > 0$ for some $t_1 > 0$. By continuity, there exist neighbourhood I_1 and I_2 of respectively z_1 and z_2 such that the infimum of p_{t_1} over $I_1 \times I_2$ is $c_1 > 0$.

Let \mathcal{K} be a compact set and let t_0 be large enough so that

$$I_1 \cap \left(\mathbb{S}^1 \times \bigcap_{(x,u) \in \mathcal{K}} [u + (\min F)t_0, u + (\max F)t_0] \right)$$

has a non-empty interior. The continuity of p_{t_0} and the compactness of \mathcal{K} imply

$$c_0 := \inf_{z \in \mathcal{K}} \mathbb{P}(Z_{t_0} \in I_1 \mid Z_0 = z) > 0.$$

Let ν be the uniform measure on I_2 , namely $\nu(D) = \frac{\lambda(D \cap I_2)}{\lambda(I_2)}$ for any Borel set D of E . Then for all $z \in \mathcal{K}$,

$$\begin{aligned} \mathbb{P}(Z_{t_0+t_1} \in D \mid Z_0 = z) &\geq \mathbb{P}(Z_{t_0+t_1} \in D \mid Z_{t_0} \in I_1) \mathbb{P}(Z_{t_0} \in I_1 \mid Z_0 = z) \\ &\geq c_0 c_1 \lambda(I_2) \nu(D). \end{aligned}$$

\square

4.2 For the velocity jump process

In this subsection we consider the process $Z = (X, Y, U)$ with generator (4). The construction of a trajectory is similar to the one exposed in Section 2.2, except from these slight modifications: in the definition of θ_1 , $g(T_k + s)$ is replaced by $U_{T_k} + \int_0^s F(X_{T_k} + uY_{T_k}) du$ and between the two jump times T_k and T_{k+1} , U is defined by $U_t = \int_{T_k}^t F(X_s) ds$.

We start with a controllability result.

Lemma 4. *Let \mathcal{K} and \mathcal{V} respectively be a compact and open set of $\mathbb{S}^1 \times \mathbb{R} \times \{-1, 1\}$. Then there exists $t_0 > 0$ such that*

$$\inf_{z \in \mathcal{K}} \mathbb{P}(Z_{t_0} \in \mathcal{V} \mid Z_0 = z) > 0.$$

Proof. The boundedness of F implies that for $t > 0$, there exists a compact set \mathcal{K}_2 such that for all $s < t$ and for all $z_0 \in \mathcal{K}$, if $Z_0 = z_0$ then $Z_s \in \mathcal{K}_2$. Hence results from [3] apply even if our whole state space is not compact. In particular, the process is Feller, and because K is compact we only need to prove that there exists t_0 such that

$$\mathbb{P}(Z_{t_0} \in \mathcal{V} \mid Z_0 = z) > 0$$

for all $z \in \mathcal{K}$. Let $z_0 = (x_0, y_0, u_0) \in \mathcal{K}$ and $z_1 = (x_1, y_1, u_1) \in \mathcal{V}$. We proceed in three steps.

First, suppose that we can deterministically choose a piecewise constant velocity $y(t) \in \{-1, 0, 1\}$, from which $(x(t), u(t))$ is defined by an initial condition and by the ODE

$$\begin{pmatrix} \dot{x} \\ \dot{u} \end{pmatrix} = \begin{pmatrix} y \\ F(x) \end{pmatrix}. \quad (11)$$

For a sufficiently large time t_0 , we can build a path between z_0 and z_1 as follows. If $u_1 > u_0$ (resp. $u_1 < u_0$), choose a non-zero velocity to bring x_0 to a point $x^* \in M(F, +)$ (resp. $x^* \in m(F, -)$). Then, pick the zero velocity and wait until u reaches the value $u_1 - \int_{x^*}^{x_1} F(s) ds$. Next, with the velocity $y = 1$, bring x^* to a point x^{**} such that $F(x^{**}) = 0$ and wait up to the time $t_0 - |x^{**} - x_1|$. Finally, with the velocity $y = 1$, push x^{**} to x_1 , and set the velocity to y_1 at time t_0 .

In a second instance, we can choose a deterministic $y(t) \in \{-1, 1\}$ such that the solution of the system (11) starting from z_0 is arbitrarily close to z_1 at time t_0 . To ensure this, we simply approximate the case $y = 0$ in the previous step by sufficiently fast and balanced jumps between -1 and 1 .

Finally, we consider the PDMP starting from z_0 . Since the random jump times have positive density, the PDMP follows arbitrarily closely a trajectory as described in the second step with positive probability. Hence, given any neighbourhood of z_1 , the PDMP has positive probability to be in it at time t_0 , which concludes. \square

Proof of Proposition 3 in the PDMP case. Consider the following vector fields:

$$G_{-1}(x, u) = \begin{pmatrix} -1 \\ F(x) \end{pmatrix} \text{ and } G_1(x, u) = \begin{pmatrix} 1 \\ F(x) \end{pmatrix}.$$

Then their difference is

$$G_1 - G_{-1} = \begin{pmatrix} 2 \\ 0 \end{pmatrix}$$

so that the Lie bracket $[G_1 - G_{-1}, G_1](x, u)$ is

$$[G_1 - G_{-1}, G_1] = 2\partial_x G_1(x, u) = \begin{pmatrix} 0 \\ 2F'(x) \end{pmatrix}$$

Since F is not constant and smooth, there exists some x such that $F'(x) \neq 0$, at which point the rank of $(G_1 - G_{-1}, [G_1 - G_{-1}, G_1])$ is 2.

According to [3, Theorem 4.4], it implies there exist a non-empty open set \mathcal{U} , a probability measure ν and $t_1, c > 0$ such that $\forall z \in \mathcal{U}$,

$$\mathbb{P}(Z_{t_1} \in \cdot \mid Z_0 = z) \geq c \nu(\cdot).$$

Thus for any $z \in \mathcal{K}$ and any Borel set D ,

$$\begin{aligned} \mathbb{P}(Z_{t_0+t_1} \in D \mid Z_0 = z) &\geq \mathbb{P}(Z_{t_0} \in \mathcal{U} \mid Z_0 = z) \times \inf_{z' \in \mathcal{U}} \mathbb{P}(Z_{t_0+t_1} \in D \mid Z_{t_0} = z') \\ &\geq \left(\inf_{z' \in \mathcal{K}} \mathbb{P}(Z_{t_0} \in \mathcal{U} \mid Z_0 = z') \right) c \nu(D) \end{aligned}$$

and Lemma 4 concludes. □

5 Proof of the main theorem

In this Section we consider either $Z = (X, U)$ or $Z = (X, U, Y)$ such as in Theorem 1, and we call E the state space, namely either $\mathbb{S}^1 \times \mathbb{R}_+$ or $\mathbb{S}^1 \times \mathbb{R}_+ \times \{-1, 1\}$.

5.1 Ergodicity when $\mathcal{M} = \emptyset$

Proof of point 1 of Theorem 1. Let $\mathcal{K} = \{z \in E, |u| \leq M\}$ where M is large enough so that Proposition 2 holds, and $\mathcal{K}' = \{z \in E, |u| \leq M + 1\}$. Let $h_0 = 0$ and

$$\begin{aligned} s_k &= \inf\{r > h_k, Z_r \in \mathcal{K}\} \\ h_{k+1} &= \inf\{r > s_k, Z_r \notin \mathcal{K}'\} \end{aligned}$$

The boundary $\partial\mathcal{K}$ being compact and the processes being Feller, the embedded Markov chain $(Z_{s_k})_{k \geq 1}$ admits an invariant measure μ_e . For a Borel set D ,

$$\mu(D) = \int \mathbb{E} \left(\int_{s_0}^{s_1} \mathbb{1}_{Z_s \in D} ds \mid Z_0 = z_0 \right) d\mu_e(z_0)$$

defines a measure which is invariant for Z (see [13, Proof of Theorem 4.1]), and has finite mass (the expectation of $s_1 - s_0$ being bounded uniformly on $z_0 \in \partial K$ from Proposition 2), which we suppose normalized to 1. Note that from the controllability results proven in Section 4, the support of μ is equal to E .

The uniqueness of the invariant measure and the exponential convergence to equilibrium are both obtained from a classical coupling argument. Let $\mu_0 = \text{Law}(Z_0)$, and let t_0 , ν and c be given by Proposition 3. Wait until the time τ_M (which according to Proposition 2 is almost surely finite, and moreover, has a finite exponential moment if so does μ_0). At time τ_M , consider two random variables Z'_{τ_M} and Θ distributed according to μ and ν respectively.

If $Z'_{\tau_M} \in \mathcal{K}$ (which happens with positive probability), we define simultaneously two processes Z and Z' with the same generator (either (3) or (4)) in such a way that $Z_{\tau_M+t_0} = \Theta = Z'_{\tau_M+t_0}$ with probability c , in which case we say the coupling is a success and from then we let Z' evolve according to its Markov dynamics (i.e. either the diffusion or the PDMP one) and set $Z_t = Z'_t$ for all $t \geq \tau_M + t_0$.

If the coupling is a failure (which happens with probability $1 - c$), we wait until Z enters \mathcal{K} again. Note that at time $\tau_M + t_0$, Z is necessarily at most at distance $t_0 \times \|F\|_\infty$ from \mathcal{K} , hence its next time of return to \mathcal{K} has some finite exponential moments that do not depend on μ_0 . Once Z has reached \mathcal{K} we try a new coupling, and so on as long as the coupling fails.

Let T be the first instance the coupling succeeds. Since μ is invariant for the dynamics, (Z, Z') is a coupling between $\text{Law}(Z_t)$ and μ , and

$$d_{TV}(\text{Law}(Z_t), \mu) \leq \mathbb{P}(T > t) \xrightarrow[t \rightarrow \infty]{} 0.$$

Uniqueness of the invariant measure is obtained by taking $\text{Law}(Z_0)$ invariant. Moreover, if μ_0 has some finite exponential moments, so does T , and the Chernoff's Inequality concludes:

$$\mathbb{P}(T > t) \leq e^{-\rho t} \mathbb{E}(e^{\rho T}).$$

□

5.2 Localization when $\mathcal{M} \neq \emptyset$

Proposition 4. *Suppose $m(F, +) \neq \emptyset$. Then there exist $p > 0$ and $M > 0$ (which does not depend on Z_0) such that if $X_0 = x_0 \in m(F, +)$ and $U_0 \geq M$, then*

$$\mathbb{P}\left(X_t \xrightarrow[t \rightarrow \infty]{} x_0\right) \geq p.$$

Proof. For $j \geq 0$, define

$$\eta_j = \frac{4 \ln(1+j)}{1+j} \wedge \delta,$$

set $c = \max\{\frac{1}{F(x)}, x \in m(F, +)\}$ and $S_0 = 0$ and define the following stopping times:

$$\begin{aligned} \tau_{j+1} &= \inf \{t > S_j, X_t \in C^{\eta_{j+1}}\}, \\ \tilde{S}_{0,j} &= S_j, \\ \tilde{T}_{k,j} &= \inf \left\{ t > \tilde{S}_{k-1,j}, X_t \in B^{\eta_{j+1}} \right\} \wedge (\tilde{S}_{k-1,j} + c) \wedge \tau_{j+1}, \quad k \geq 1, \\ \tilde{S}_{k,j} &= \inf \left\{ t > \tilde{T}_{k,j}, X_t \in A \right\} \wedge \tau_{j+1}, \quad k \geq 1. \end{aligned}$$

Let

$$N_j = \inf \left\{ k \in \mathbb{N}, \tilde{S}_{k,j} \geq S_j + c \text{ or } \tilde{S}_{k,j} = \tau_{j+1} \right\}$$

and $S_{j+1} = \tilde{S}_{N_j,j}$.

Let us give some intuition on these definitions. The connected component of $(C^{\eta_j})^c$ that contains x_0 is a neighbourhood of x_0 whose diameter goes to 0 as j goes to ∞ . At time τ_j , the process has escaped from this neighbourhood. For $t \leq \tau_j$, the process makes possibly many oscillations near x_0 . When such an oscillation is large enough for the process to reach B^{η_j} (this is at a time $\tilde{T}_{k,j}$ for some k), we consider this is the beginning of an attempt to leave $(C^{\eta_j})^c$. If this attempt fails, the process falls back to x_0 (this is $\tilde{S}_{k,j}$). While X makes those attempts to escape, time goes by, so that U increases: after a time c , U has increased at least by 1. Next time X falls back to x_0 (this is S_{j+1}), we shrink the neighbourhood, namely from then we consider that the process escapes if it reaches $C^{\eta_{j+1}}$. From S_j to S_{j+1} , there have been N_j attempts to leave. The sequence η is scaled so that there is in fact a positive probability that the process never escape from the shrinking neighbourhood that collapses at infinity to $\{x_0\}$.

Let us write these ideas more precisely. Note that as long as $S_{j+1} < \tau_{j+1}$,

$$S_{j+1} - S_j \geq c \quad \text{and} \quad U_t \geq M + j$$

for $t \geq S_j$. We take M large enough so that $(M + j)\eta_j > 1$ for all $j \in \mathbb{N}$. Therefore, from Proposition 1, for all $k \geq 1$,

$$\mathbb{P}(\tilde{S}_{k,j} = \tau_{j+1} | \tilde{T}_{k,j} < \tau_{j+1}) \leq K(j + M)e^{-(j+M)\eta_j}.$$

It implies that $\left(\mathbb{1}_{\tilde{S}_{(i \wedge N_j),j} < \tau_{j+1}} + (i \wedge N_j)K(j + M)e^{-(j+M)\eta_j} \right)_{i \geq 0}$ is a submartingale. Thus,

$$\begin{aligned} \mathbb{P}(S_{j+1} < \tau_{j+1} | S_j < \tau_j) &= 1 + \mathbb{E}(\mathbb{1}_{S_{j+1} < \tau_{j+1}} - \mathbb{1}_{S_j < \tau_j} | S_j < \tau_j) \\ &\geq 1 - K(j + M)e^{-(j+M)\eta_{j+1}}\mathbb{E}(N_j | S_j < \tau_j). \end{aligned} \quad (12)$$

From Proposition 1, we have

$$\tilde{S}_{k+1,j} - \tilde{S}_{k,j} \stackrel{sto}{\geq} \sqrt{\eta_j}R.$$

Hence, considering a sequence $(R_i)_{i \in \mathbb{N}}$ of i.i.d random variables distributed like R

$$\begin{aligned} N_j &\stackrel{sto}{\leq} \inf \left\{ n \geq 1, \sqrt{\eta_j} \sum_{i=1}^n R_i \geq c \right\} \\ &\leq \left\lceil \frac{2c}{\mathbb{E}(R_1)\sqrt{\eta_j}} \right\rceil + \inf \left\{ n \geq 1, \frac{1}{n} \sum_{i=1}^n R_i \geq \frac{\mathbb{E}(R_1)}{2} \right\}. \end{aligned}$$

Because R satisfies a Large Deviation Principle,

$$\begin{aligned} \mathbb{E}(N_j | S_j < \tau_j) &\leq \left\lceil \frac{2c}{\mathbb{E}(R_1)\sqrt{\eta_j}} \right\rceil + \sum_{n \geq 1} \mathbb{P} \left(\frac{1}{n} \sum_{i=1}^n R_i \leq \frac{\mathbb{E}(R_1)}{2} \right) \\ &\leq \frac{K'}{\sqrt{\eta_j}} \end{aligned}$$

for some constant K' which does not depend on j , nor M . Thus (12) is now

$$\mathbb{P}(S_{j+1} < \tau_{j+1} | S_j < \tau_j) \geq 1 - \frac{K'K}{\sqrt{\eta_j}}(j+M)e^{-(j+M)\eta_{j+1}}.$$

Take M large enough so that the right-hand side is positive for all $j \in \mathbb{N}$. Then by induction

$$\begin{aligned} \mathbb{P}(S_{j+1} < \tau_{j+1}) &= \mathbb{P}(S_{j+1} < \tau_{j+1} | S_j < \tau_j) \mathbb{P}(S_j < \tau_j) \\ &\geq \prod_{i=0}^{j+1} \left(1 - \frac{K'K}{\sqrt{\eta_i}}(i+M)e^{-(i+M)\eta_{i+1}} \right). \end{aligned}$$

As $(\{S_j < \tau_j\})_{j \geq 1}$ is a decreasing family of events,

$$\begin{aligned} \mathbb{P}(S_j < \tau_j \forall j \in \mathbb{N}) &= \lim_{j \rightarrow \infty} \mathbb{P}(S_j < \tau_j) \\ &\geq \prod_{j=0}^{\infty} \left(1 - \frac{K'K}{\sqrt{\eta_j}}(j+M)e^{-(j+M)\eta_{j+1}} \right) \\ &= \exp \left(\sum_{j \geq 0} \ln \left(1 - \frac{K'K}{\sqrt{\eta_j}}(j+M)e^{-(j+M)\eta_{j+1}} \right) \right). \end{aligned}$$

For j large enough,

$$\frac{K'K}{\sqrt{\eta_j}}(j+M)e^{-(j+M)\eta_{j+1}} \leq \frac{1}{j^2}$$

and

$$\ln \left(1 - \frac{K'K}{\sqrt{\eta_j}}(j+M)e^{-(j+M)\eta_{j+1}} \right) \geq -\frac{1}{2j^2}$$

which means $\mathbb{P}(S_j < \tau_j \forall j \in \mathbb{N}) > p > 0$ where p does not depend on Z_0 . Yet,

$$\{S_j < \tau_j \forall j \in \mathbb{N}\} = \{\forall j \in \mathbb{N}, \forall s \geq S_j, X_s \in I_{x_0}^{\eta_j}\}$$

and the S_j 's are all a.s. finite, which concludes. \square

Remark 4. The proof even gives a speed of convergence. Indeed we can see that $S_{j+1} \stackrel{sto}{\leq} S_j + c + \sqrt{\delta}R$, so that the S_j 's grow linearly to infinity. From the non-degeneracy assumption on F , there exist $n \in \mathbb{N}$ and $c > 0$ such that the diameter of $I_{x_0}^{\eta_j}$ is less than $c\eta_j^{\frac{1}{n}}$, depending on the first derivative of F at x_0 to be non-zero (if F is a Morse function, $n = 2$). It means when there is convergence, it occurs at least at a speed of order $(\frac{\ln t}{t})^{\frac{1}{n}}$.

Proof of point 2 of Theorem 1: First note that by changing U and F to their opposites, Proposition 4 also says that if $M(F, -) \neq \emptyset$ then there exist $p, M > 0$ such that if $U_0 < -M$ and $X_0 \in M(F, -)$ then X_t converges to x_0 with probability at least p .

For $M > 0$ large enough, $\varepsilon > 0$ small enough and $x \in \mathcal{M}$, let

$$\begin{aligned}\mathcal{V}_x^\varepsilon &= \{z' \in E \text{ s.t. } |x' - x| < \varepsilon \text{ and } u' \times \text{sign}(F(x)) > M\} \\ \mathcal{V}^\varepsilon &= \bigcup_{x \in \mathcal{M}} \mathcal{V}_x^\varepsilon.\end{aligned}$$

When ε is fixed, for M large enough, if the process starts in $\mathcal{V}_x^\varepsilon$, from Inequality (7) (which is written for $x \in m(F, +)$ but by symmetry, again, also holds for $x \in M(F, -)$) it has a probability at least $\frac{1}{2}$ to hit \mathcal{V}_x^0 before leaving $\mathcal{V}_x^{2\varepsilon}$. Then from Proposition 4, X has a probability at least p to converge to x .

Let

$$\mathcal{K} = \{z \in E, |u| \leq M\}.$$

As long as the process is in the complementary of $\mathcal{K} \cup \mathcal{V}^\varepsilon$, the conclusion of Proposition 2 (together with Remark 3) holds so that, denoting by τ_D the first hitting time of a set D , we have

$$\mathbb{P}(\tau_{\mathcal{V}^\varepsilon} \wedge \tau_{\mathcal{K}} < \infty | Z_0 = z) = 1.$$

for all $z \in E$ (more precisely: suppose $U_0 > M$, the case $U_0 < -M$ being symmetric. Then we can define a potential \tilde{F} which is equal to F away from $m(F, +)$ and which have no positive minimum, from which we can define an associated process \tilde{Z} with the initial condition $\tilde{Z}_0 = Z_0$ such that $\tilde{Z}_t = Z_t$ as long as $Z \notin \mathcal{K} \cup \mathcal{V}^\varepsilon$. Then \tilde{Z} , to which Proposition 2 applies, hits \mathcal{K} in a finite time).

On the other hand, by Lemmas 4 (for the PDMP) and 3 (for the diffusion), there exists $t_0 > 0$ such that for all $x \in \mathcal{M}$,

$$\inf_{z \in \mathcal{K}} \mathbb{P}(Z_{t_0} \in \mathcal{V}_x^\varepsilon | Z_0 = z) > 0.$$

It therefore follows that for any $z \in E$,

$$\mathbb{P}(\tau_{\mathcal{V}^0} < \infty | Z_0 = z) = 1$$

and moreover

$$\mathbb{P}(X_{\tau_{\mathcal{V}^0}} = x) > 0$$

for all $x \in \mathcal{M}$. Proposition 4 concludes. \square

6 Acknowledgement

The first author thanks Olivier Raimond for a discussion that led to this paper. Finally, we acknowledge financial support from the Swiss National Science Foundation, Grant 200021_163072.

References

- [1] Benaïm M. and Gauthier C.E.: *Self-repelling diffusions on a Riemannian manifold*, Probab. Theory Relat. Fields., Article first published online: 06 May 2016, DOI: 10.1007/s00440-016-0717-1 (2016)
- [2] Benaïm M.: Ledoux M. and Raimond O., *Self-interacting diffusions*, Probab. Theory Related Fields, volume 122, number 1: 1-41 (2002)
- [3] Benaïm M.: Le Borgne S., Malrieu F. and Zitt P.-A., *Qualitative properties of certain piecewise deterministic Markov processes*, Annales de l'Institut Henri Poincaré (Probabilités et Statistiques), 51(3): 1040-1075 (2015)
- [4] Bovier, A., Eckhoff, M., Gaynard, V. and Klein, M.: *Metastability in reversible diffusion processes. I. Sharp asymptotics for capacities and exit times*, J. Eur. Math. Soc., Volume 6, 399–424 (2004)
- [5] Ciesielski Z. and Taylor S.J.: *First passage times and sojourn times for Brownian motion in space and the exact Hausdorff measure of the sample path*, Trans. Amer. Math. Soc. 103: 434-450 (1962).
- [6] Cranston M. and Le Jan Y.: *Self attracting diffusions : Two case studies*, Math. Ann. 303(1): 87-93 (1995)
- [7] Fontbona J., Guerin H. and Malrieu F.: *Long time behavior of telegraph processes under convex potentials*, Stoch. Proc. & Applic. Article first published online: 19 April 2016, DOI: 10.1016/j.spa.2016.04.002 (2016)
- [8] Gauthier C-E.: *Self attracting diffusions on a sphere and application to a periodic case*, arXiv preprint 1501.04827 (2015)
- [9] Hairer, M.: *On Malliavin's proof of Hörmander's theorem*, Bull. Sci. Math 135 (6-7): 650-666 (2011)
- [10] Herrmann S. and Roynette B.: *Boundedness and convergence of some self-attracting diffusions*, Math. Ann. 325(1): 81-96 (2003)
- [11] Ikeda, N., Watanabe, S.: *Stochastic Differential Equations and Diffusion Processes*, North Holland, New-York, 1981
- [12] Karatzas, I., Shreve, S.E.: *Brownian motion and stochastic calculus*, Springer-Verlag, New-York, second edition, 1991
- [13] Khasminskii, R.: *Stochastic stability of Differential Equations*, Second Edition, Stochastic Modelling and Applied Probability, Springer, volume 66 (2012).
- [14] Monmarché, P. *Piecewise deterministic simulated annealing*. ALEA, Lat. Am. J. Probab. Math. Stat. 13 (1), 357-398 (2016)

- [15] Rassoul-Agha, F., Seppäläinen, T.: *A Course on Large Deviations with an Introduction to Gibbs Measures*, American Mathematical Society, Graduate Studies in Mathematics, Volume 162 (2015)
- [16] Royer G., *A remark on simulated annealing of diffusion processes*, SIAM J. Control Optim., Volume 27, number 6, 1403-1408 (1989)